

1 **ABSTRACT**

2 Data classification using stochastic key feature generation includes
3 obtaining a set of training data having associated summaries. The set of training
4 data and associated summaries are used to generate a key feature generation
5 model. Another set of training data having associated categories is also obtained,
6 and the key feature generation model is used to map this other set of training data
7 to a set of vectors. A data classifier is then trained based on the set of vectors and
8 the associated categories.

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25